

# Riemann integration via primitives for a new proof to the change of variable theorem

May, 2011

Dedicated to Hyman Kestelman and Roy O. Davies on the 50th  
anniversary of the publication of their excellent papers on change of  
variable.

Rodrigo López Pouso  
Departamento de Análise Matemática  
Facultade de Matemáticas,  
Universidade de Santiago de Compostela, Campus Sur  
15782 Santiago de Compostela, Spain.

## Abstract

We approach the Riemann integral via generalized primitives to give a new proof for a general result on change of variable originally proven by Kestelman and Davies. Our proof is similar to Kestelman's, but we hope readers will find it clearer thanks to the use of a new test for the Riemann integrability (which we introduce in this paper) along with some ingredients from some other more recent proofs available in the literature. We also include a bibliographical review of related results and proofs. Although this paper emphasizes in the change of variable theorem, our contributions to the Riemann integration theory are of independent interest. For instance, we present a very simple proof to the fact that continuous functions are integrable which avoids the use of uniform continuity.

# 1 Introduction

Let  $G : I = [a, b] \rightarrow \mathbb{R}$  have a continuous derivative on the interval  $I$ , and let  $f : G(I) \rightarrow \mathbb{R}$  be continuous on  $G(I)$ .

If  $G$  is not constant then we can use the chain rule and the Fundamental Theorem of Calculus to compute

$$\left( \int_{G(a)}^{G(t)} f(x) dx \right)' = f(G(t)) G'(t) \quad (t \in [a, b]),$$

and now it suffices to integrate between  $a$  and  $b$  to finish the nontrivial part of the proof of the following well-known theorem for the Riemann integral:

**Theorem 1.1** *The change of variables formula*

$$\int_{G(a)}^{G(b)} f(x) dx = \int_a^b f(G(t)) G'(t) dt \quad (1.1)$$

is valid provided that  $G : I = [a, b] \rightarrow \mathbb{R}$  is continuously differentiable on  $I$  and  $f : G(I) \rightarrow \mathbb{R}$  is continuous on  $G(I)$ .

Theorem 1.1 plays a fundamental role in elementary courses on Riemann integration: it is the key for evaluating exactly thousands of integrals. Moreover, Theorem 1.1 is stronger than it might appear at first glance. Notice that the substitution  $G$  needs not be differentiable everywhere: it suffices that we can split the right-hand side in (1.1) as a finite sum of integrals each of which satisfies the conditions of Theorem 1.1. In fact, as we will show, much more general results are known in the theory of Riemann integration.

Ch. J. de la Vallée Poussin established a change of variable formula for Lebesgue integrals at the beginning of the 20th century, see [22]. The availability of that result reduced later interests of the mathematical community in similar ones with Riemann integrals. Therefore, it is not a surprise that we have to wait until 1961 for the following general theorem, when it was proven by H. Kestelman [5] and R. O. Davies [3] in two consecutive papers in the same journal. From now on, and unless stated otherwise, integrability is to be understood in the Riemann sense.

**Theorem 1.2 (Kestelman's Theorem)** *Assume that  $g : I = [a, b] \rightarrow \mathbb{R}$  is integrable on  $I$  and let  $G(t) = c + \int_a^t g(s) ds$  for all  $t \in I$  and some  $c \in \mathbb{R}$ .*

*If  $f : G(I) \rightarrow \mathbb{R}$  is integrable on  $G(I)$  then  $(f \circ G)g$  is integrable on  $I$  and*

$$\int_{G(a)}^{G(b)} f(x) dx = \int_a^b f(G(t))g(t) dt.$$

Kestelman proved Theorem 1.2 first, and then Davies found an elementary proof which, in particular, avoids the concept of null measure set, involved in Kestelman's original proof. Notice that, as pointed out by Kestelman, the assumptions in Theorem 1.2 imply that the change of variable formula is valid with Lebesgue integrals (see, for instance, [8, 14] or Theorem 6.95 on page 325 in [16]). However, change of variable theorems for Lebesgue integrals give no information about the integrability of  $(f \circ G)g$  in the Riemann sense, and therefore Kestelman's Theorem is neither a particular case to any of them nor a piece of completely romantic mathematics.

Kestelman's Theorem and Davies' elementary proof have been revised and rediscovered several times, see [1, 2, 7, 10, 11, 12, 13, 19, 23] and our last section on bibliographical reviews. In this introduction we only point out the work done by Preiss and Uher in [12], who established the following more general version of Theorem 1.2:

**Theorem 1.3** *Let  $g : I = [a, b] \rightarrow \mathbb{R}$  be integrable on  $I$ ,  $c \in \mathbb{R}$ , and  $G(t) = c + \int_a^t g(s) ds$  for all  $t \in I$ .*

*If  $f : G(I) \rightarrow \mathbb{R}$  is bounded on  $G(I)$  then  $f$  is integrable on  $G(I)$  if and only if  $(f \circ G)g$  is integrable on  $I$  and, in that case, we have*

$$\int_{G(a)}^{G(b)} f(x) dx = \int_a^b f(G(t))g(t) dt. \quad (1.2)$$

In this paper we present a new proof to Theorem 1.3 which keeps within the theory of Riemann integration and which we hope that readers may find more readable than others. Roughly, we adapt the classical proof described at the beginning of the introduction, in the sense we use the fact that the mapping

$$t \mapsto \int_{G(a)}^{G(t)} f(x) dx$$

is a “generalized” primitive of  $(f \circ G)g$ .

## 2 Riemann integrability through primitives

Teachers always warn students not to make the following typical mistake: the definition of the Riemann integral of a function  $f : [a, b] \rightarrow \mathbb{R}$  is

$$\int_a^b f(x) dx = F(b) - F(a), \quad \text{where } F \text{ is a primitive of } f \text{ } (F' = f \text{ in } [a, b]).$$

As we all know, integrable functions need not have primitives at all, so the previous formula is not adequate as a definition.

It is however true that we can define the Riemann integral by means of primitives, see [18], although not primitives in the usual sense. In this section we propose an easy way to do it and we highlight some interesting consequences.

We are going to use the most basic concepts from the Riemann integration theory, which can be looked up in textbooks such as [15, 21].

To motivate our new definition of primitive, let  $f : [a, b] \rightarrow \mathbb{R}$  be bounded and consider the indefinite lower integral

$$F_*(x) = \int_a^x f(y) dy \quad (x \in [a, b]). \quad (2.3)$$

For  $x, y \in [a, b]$ ,  $x < y$ , we have (thanks to the additivity of the lower integral with respect to the intervals of integration)

$$F_*(y) - F_*(x) = \int_x^y f(z) dz,$$

and, therefore,

$$\inf_{x \leq z \leq y} f(z) \leq \frac{F_*(y) - F_*(x)}{y - x} \leq \sup_{x \leq z \leq y} f(z).$$

Notice that the same property is fulfilled by the function

$$F^*(x) = \int_a^x f(y) dy \quad (x \in [a, b]). \quad (2.4)$$

We have just shown that bounded functions have “generalized primitives” according to the following definition (which does not involve derivatives).

**Definition 2.1** *Let  $f : [a, b] \rightarrow \mathbb{R}$  be a bounded function.*

*A generalized primitive of  $f$  is a function  $F : [a, b] \rightarrow \mathbb{R}$  such that for all  $x, y \in [a, b]$ ,  $x < y$ , we have*

$$\inf_{x \leq z \leq y} f(z) \leq \frac{F(y) - F(x)}{y - x} \leq \sup_{x \leq z \leq y} f(z). \quad (2.5)$$

Generalized primitives need not be differentiable, but (2.5) readily yields some relations between their Dini derivatives and the function  $f$ .

**Proposition 2.1** *Let  $f : [a, b] \longrightarrow \mathbb{R}$  be bounded and let  $F$  be a generalized primitive of  $f$ .*

*For every  $x \in [a, b)$  we have*

$$\begin{aligned} \min \left\{ f(x), \liminf_{y \rightarrow x^+} f(y) \right\} &\leq D_+ f(x) \\ &\leq D^+ f(x) \leq \max \left\{ f(x), \limsup_{y \rightarrow x^+} f(y) \right\}, \end{aligned} \quad (2.6)$$

*and for every  $x \in (a, b]$  we have*

$$\begin{aligned} \min \left\{ f(x), \liminf_{y \rightarrow x^-} f(y) \right\} &\leq D_- f(x) \\ &\leq D^- f(x) \leq \max \left\{ f(x), \limsup_{y \rightarrow x^-} f(y) \right\}. \end{aligned} \quad (2.7)$$

*In particular, if  $f$  is continuous at  $x \in [a, b]$  then  $F'(x) = f(x)$  (as usual, by  $F'(a)$  we mean the right-hand derivative, and by  $F'(b)$  the left-hand derivative).*

**Proof.** We will only prove the left-hand inequality in (2.6) because the remaining inequalities can be proven in analogous ways. Let  $x \in [a, b)$  be fixed; we simply take limit inferior in the first inequality of (2.5) and we get the desired result:

$$\begin{aligned} D_+ f(x) &= \liminf_{y \rightarrow x^+} \frac{F(y) - F(x)}{y - x} \geq \liminf_{y \rightarrow x^+} \inf_{x \leq z \leq y} f(z) = \lim_{y \rightarrow x^+} \inf_{x \leq z \leq y} f(z) \\ &= \lim_{y \rightarrow x^+} \min \left\{ f(x), \inf_{x < z \leq y} f(z) \right\} \\ &= \min \left\{ f(x), \liminf_{y \rightarrow x^+} f(y) \right\}. \end{aligned}$$

The existence of derivatives when  $f$  is continuous is an immediate consequence of (2.6) and (2.7).  $\square$

When  $f$  is continuous generalized primitives are exactly the usual ones.

**Corollary 2.2** *If  $f : [a, b] \longrightarrow \mathbb{R}$  is continuous in  $[a, b]$  then every generalized primitive of  $f$  is a primitive in usual sense and, conversely, primitives are generalized primitives.*

**Proof.** Proposition 2.1 ensures that generalized primitives satisfy  $F' = f$  in  $[a, b]$ , so they are primitives in the usual sense.

Conversely, if  $F$  is a primitive of  $f$  then for every  $x, y \in [a, b]$ ,  $x < y$ , the Mean Value Theorem guarantees the existence of some  $z \in (x, y)$  such that

$$\frac{F(y) - F(x)}{y - x} = F'(z) = f(z),$$

which implies (2.5).  $\square$

Now we are in a position to establish our first test for the Riemann integrability in terms of generalized primitives. The Barrow's Rule for computing integrals using generalized primitives is included in the test.

**Theorem 2.3** *Let  $f : [a, b] \rightarrow \mathbb{R}$  be bounded and  $A \in \mathbb{R}$ . The following two statements are equivalent:*

1. *The function  $f$  is integrable on  $[a, b]$  and  $\int_a^b f(x) dx = A$ ;*
2. *Every generalized primitive of  $f$  satisfies  $F(b) - F(a) = A$ .*

**Proof.** Let  $F : [a, b] \rightarrow \mathbb{R}$  be an arbitrary generalized primitive of  $f$  and let  $P = \{x_0, x_1, \dots, x_n\}$  be a partition of  $[a, b]$ .

By (2.5), we have

$$F(b) - F(a) = \sum_{k=1}^n [F(x_k) - F(x_{k-1})] \leq U(f, P),$$

and, similarly,  $F(b) - F(a) \geq L(f, P)$ . Since  $P$  was arbitrary and  $f$  is integrable on  $[a, b]$ , we conclude that  $F(b) - F(a) = \int_a^b f(x) dx = A$ .

Conversely, we take into account that  $F_*$  and  $F^*$ , as defined in (2.3) and (2.4), are generalized primitives of  $f$ . Hence condition 2 ensures that

$$\int_a^b f(s) ds = F_*(b) - F_*(a) = A = F^*(b) - F^*(a) = \overline{\int_a^b f(s) ds}.$$

$\square$

Forgetting about computations, we have the following concise test for the Riemann integrability.

**Theorem 2.4** *A bounded function  $f : [a, b] \rightarrow \mathbb{R}$  is integrable if and only if any pair of its generalized primitives differ in a constant.*

**Proof.** Assume that  $f$  is integrable and let  $F$  and  $G$  be two of its generalized primitives. For each  $x \in (a, b]$  we use Theorem 2.3 in the interval  $[a, x]$  to deduce that  $F(x) - F(a) = G(x) - G(a)$  or, equivalently, that  $F(x) - G(x) = F(a) - G(a)$ .

Conversely, if  $F$  and  $G$  are two arbitrary primitives of  $f$  the assumption ensures the existence of some constant  $c \in \mathbb{R}$  such that  $F(x) = c + G(x)$  for all  $x \in [a, b]$ . Hence  $F(b) - F(a) = G(b) - G(a)$ , which implies that  $f$  is integrable on  $[a, b]$  by Theorem 2.3.  $\square$

When  $f$  is integrable then its generalized primitives are exactly its indefinite integrals. More precisely we have the following corollary.

**Corollary 2.5** *If  $f : [a, b] \rightarrow \mathbb{R}$  is integrable then  $F : [a, b] \rightarrow \mathbb{R}$  is a generalized primitive of  $f$  if and only if for each  $x_0 \in [a, b]$  there exists some  $c \in \mathbb{R}$  such that*

$$F(x) = c + \int_{x_0}^x f(y) dy \quad (x \in [a, b]).$$

**Proof.** Notice that the indefinite integral

$$x \in [a, b] \mapsto \int_{x_0}^x f(y) dy$$

is a generalized primitive of  $f$  which, according to Theorem 2.4, is equal to any other generalized primitive up to adding some constant.  $\square$

**Remark 2.1** *In [20, Theorem 1] Thomson solves completely the question of which functions  $F$  are indefinite integrals of some unknown Riemann-integrable function  $f$ . Corollary 2.5 reveals that indefinite integrals are exactly generalized primitives, so we conclude, thanks to [20, Theorem 1], that  $F$  is a generalized primitive of some unknown integrable function if and only if  $F$  satisfies condition (3) in [20], namely, if for all  $\varepsilon > 0$  a positive  $\delta$  can be found so that*

$$\sum_{i=1}^n \left| \frac{F(\xi_i) - F(x_{i-1})}{\xi_i - x_{i-1}} - \frac{F(x_i) - F(\xi'_i)}{x_i - \xi'_i} \right| (x_i - x_{i-1}) < \varepsilon$$

*for every subdivision  $a = x_0 < x_1 < \dots < x_n = b$  that is finer than  $\delta$  and every choice of associated points  $x_{i-1} < \xi_i \leq \xi'_i < x_i$ .*

We are going to illustrate the applicability of Theorem 2.4 in the following two well-known situations.

**Example 2.1** *Continuous functions are integrable, and the usual way to prove it uses uniform continuity. Theorem 2.4 yields an alternative and very easy proof. In particular, we avoid the uniform continuity (perhaps the following proof alone justifies the introduction of generalized primitives).*

**Proof.** Let  $f : [a, b] \rightarrow \mathbb{R}$  be continuous on  $[a, b]$  and let  $F$  and  $G$  be two of its generalized primitives. According to Corollary 2.2,  $F$  and  $G$  are primitives in the usual sense, so we have  $(F - G)' = 0$  in  $[a, b]$  and then  $F - G$  is constant. We conclude that  $f$  is integrable on  $[a, b]$  by virtue of Theorem 2.4.  $\square$

Our second example concerns a typical non-integrable bounded function.

**Example 2.2** *Let  $f(x) = 0$  for  $x \in [0, 1] \cap \mathbb{Q}$  and  $f(x) = 1$  for  $x \in [0, 1] \setminus \mathbb{Q}$ . The functions*

$$F_\lambda(x) = \lambda x \quad (x \in [0, 1]),$$

*are generalized primitives of  $f$  provided that  $\lambda \in [0, 1]$ . Obviously,  $F_1$  and  $F_0$  do not differ in a constant, hence  $f$  is not integrable on  $[0, 1]$ .*

Using Theorem 2.4 to deduce integrability needs checking that generalized primitives differ in a constant. The following lemma is very useful for that task, and it will be fundamental in our proof of Theorem 1.3 in the next section. For different proofs see [6, Proposition 1], [13, Theorem 2] or [16, Lemma 6.89].

**Lemma 2.6** *If  $F$  is Lipschitz on  $[a, b]$  and  $F'(x) = 0$  for almost all  $x \in [a, b]$  then  $F$  is constant on  $[a, b]$ .*

The combination of Lemma 2.6 and Theorem 2.4 is powerful, and we show it by proving the sufficient part of Lebesgue's test for Riemann-integrability, which ensures that integrable functions are exactly those bounded functions whose sets of discontinuity points have zero measure. Here is the precise statement and its short proof.

**Proposition 2.7** *If  $f : [a, b] \rightarrow \mathbb{R}$  is bounded and continuous almost everywhere in  $[a, b]$  then  $f$  is integrable on  $[a, b]$ .*

**Proof.** Let  $F$  and  $G$  be two generalized primitives of  $f$ . Then  $F - G$  is a Lipschitz continuous function satisfying  $(F - G)' = 0$  almost everywhere. Lemma 2.6 guarantees now that  $F - G$  is constant.  $\square$



Finally, note that Lebesgue's test and Proposition 2.1 guarantee that if  $f$  is integrable and  $F$  is one of its generalized primitives, then  $F'(x) = f(x)$  for almost all  $x \in [a, b]$ . Conversely, we have the following result.

**Proposition 2.8** *If  $f : [a, b] \rightarrow \mathbb{R}$  is bounded,  $F : [a, b] \rightarrow \mathbb{R}$  is absolutely continuous, and  $F'(x) = f(x)$  a.e. in  $[a, b]$ , then  $F$  is a generalized primitive of  $f$ .*

**Proof.** (This proof is the only part in this paper where integrals are to be understood in the Lebesgue sense. Notice that we will not use Proposition 2.8 in the next section.) For  $x, y \in [a, b]$ ,  $x < y$ , the Fundamental Theorem of Calculus for the Lebesgue integral yields

$$F(y) - F(x) = \int_x^y F'(z) dz = \int_x^y f(z) dz,$$

which implies (2.5). □

**Remark 2.2** *We have shown a way to introduce the Riemann integral from generalized primitives immediately after studying lower and upper Darboux integrals and we saw that it has some advantages.*

*We point out now another advantage: the definition of generalized primitive can be adapted for other intervals than compact ones, and it could therefore be possible to unify the proper and improper theories of Riemann integration thanks to generalized primitives. We will not go further with this idea, because similar and deeper approaches have already been successfully tried. For instance, Talvila [18] defines integrals by means of primitives in the weak or distributional sense (to which the adjective “generalized” would fit better than to ours).*

### 3 Change of variables

This section is devoted to proving Theorem 1.3 following the ideas in [5, 13] combined with our theorems 2.3 and 2.4. The proof due to Sarkhel and Výborný in [13] is similar to Kestelman's and it consists of two steps, proving, loosely speaking, Theorem 1.3 for  $G$  monotone first and using it to establish the general case. Here we work directly with the general case.

We also use in our proof the well-known result which guarantees that Lipschitz continuous functions map null-measure sets into null-measure sets. See [5, Lemma 1] or [16, Lemma 6.87].

**Proof of Theorem 1.3.** Assume that  $f$  is integrable on  $G(I)$ , let  $H = H(t)$  be an arbitrary generalized primitive of  $(f \circ G)g$  and let

$$\mathcal{F}(t) = \int_{G(a)}^{G(t)} f(x) dx \quad (t \in [a, b]).$$

If we prove that  $\mathcal{F} - H$  is a constant then we can conclude that every generalized primitive of  $(f \circ G)g$  is of the form  $H = c + \mathcal{F}$  for some constant  $c \in \mathbb{R}$ , and therefore, Theorem 2.3 guarantees that  $(f \circ G)g$  is integrable and

$$\int_a^b f(G(t))g(t) dt = \mathcal{F}(b) - \mathcal{F}(a) = \int_{G(a)}^{G(b)} f(x) dx.$$

To prove that  $\mathcal{F} - H$  is constant we note that  $\mathcal{F} - H$  is Lipschitz and we use Lemma 2.6, so it suffices to prove that  $(\mathcal{F} - H)' = 0$  almost everywhere in  $(a, b)$ . To do it, we first notice that we can neglect all those points  $t \in (a, b)$  where  $g$  is discontinuous, because they form a null-measure set. Therefore we have only to study the subset of  $(a, b)$  where  $g$  is continuous. We split this set into

$$A = \{t \in (a, b) : g \text{ continuous at } t \text{ and } g(t) = 0\}$$

and

$$B = \{t \in (a, b) : g \text{ continuous at } t \text{ and } g(t) \neq 0\}.$$

For all  $t \in A$  the function  $(f \circ G)g$  is continuous at  $t$ , so Proposition 2.1 ensures that  $H'(t) = f(G(t))g(t) = 0$ . In turn, for  $t \in A$  and  $h \neq 0$ ,  $|h|$  sufficiently small, we have

$$\left| \frac{\mathcal{F}(t+h) - \mathcal{F}(t)}{h} \right| \leq \left| \frac{\int_t^{t+h} g(s) ds}{h} \right| \sup_{x \in G(I)} |f(x)|,$$

and then  $\mathcal{F}'(t) = 0$ . Hence  $(\mathcal{F} - H)' = 0$  everywhere in  $A$ .

We turn our attention now to the set  $B$ , which we first decompose as  $B = C \cup D$ , where

$$C = \{t \in B : f \text{ is continuous at } G(t)\} \text{ and } D = B \setminus C.$$

For all  $t \in C$  the function  $(f \circ G)g$  is continuous at  $t$ , so Proposition 2.1 ensures  $H'(t) = f(G(t))g(t)$ , and the chain rule gives  $\mathcal{F}'(t) = f(G(t))g(t)$ . We have proven that  $(\mathcal{F} - H)' = 0$  everywhere in  $C$ .

Finally, we prove that

$$D = \{t \in (a, b) : g \text{ continuous at } t, g(t) \neq 0, f \text{ discontinuous at } G(t)\}$$

is a null-measure set. To do it we mimic some ideas in Serrin and Varberg's proof of [14, Theorem 1] and in Kestelman's proof of [5, Lemma 2].<sup>1</sup>

For each  $t \in D$  we can find a sufficiently large  $n \in \mathbb{N}$  such that  $|g| > 1/n$  and  $g$  does not change sign in  $[t - 1/n, t + 1/n]$ , so  $D$  is expressible as a countable union of sets of the form

$$D_\mu = \{t \in D : |g| > \mu \text{ and } g \text{ has constant sign in } [t - \mu, t + \mu] \cap I\} \quad (\mu > 0).$$

Now we prove that for each  $\mu > 0$  the set  $D_\mu$  is a null-measure set, and it suffices to show that  $D_\mu \cap J$  is a null-measure set for an arbitrary interval  $J \subset I$  with length less than  $\mu/2$ .

If  $D_\mu \cap J$  is empty we are done, so assume that there is some  $t \in D_\mu \cap J$ . The definition of  $D_\mu$  and the length of  $J$  ensure that  $|g| > \mu$  and  $g$  does not change sign in  $J$ . Hence the restriction  $G|_J$  is strictly monotone and it has an inverse  $(G|_J)^{-1}$  which is Lipschitz continuous (with Lipschitz constant  $1/\mu$ ). Now we have  $D_\mu \cap J = (G|_J)^{-1}(G(D_\mu \cap J))$ , a null-measure set because  $G(D_\mu \cap J)$  is contained in the set of discontinuity points of  $f$  and  $(G|_J)^{-1}$  is Lipschitz continuous. The proof that  $D$  is a null-measure set is finished.

Hence  $(\mathcal{F} - H)' = 0$  a.e. in  $[a, b]$ , and the first part of the proof is over.

Conversely, assume that  $(f \circ G)g$  is integrable and define

$$\mathcal{H}(t) = \int_a^t f(G(s))g(s) ds \quad (t \in [a, b]).$$

Now let  $F$  be a generalized primitive of  $f$  in the interval  $G(I)$ , and consider the composition  $\mathcal{F} = F \circ G$ .

Adjusting the previous arguments one can show that  $(\mathcal{F} - \mathcal{H})' = 0$  a.e. in  $[a, b]$  (Proving that  $D$  is a null-measure set is easier in this case: it suffices to note that  $D$  is contained in the set of discontinuity points of  $(f \circ G)g$ , which is a null-measure set<sup>2</sup>). Hence there is some  $c \in \mathbb{R}$  such that

$$\mathcal{F}(t) = F(G(t)) = c + \mathcal{H}(t) \quad \text{for all } t \in I, \quad (3.8)$$

---

<sup>1</sup>We are more explicit on this in our review of Kestelman's paper in the next section.

<sup>2</sup>It seems that Kestelman and Davies dealt with the hardest part in Theorem 1.3.

and, in particular,

$$\int_a^b f(G(s))g(s) ds = \mathcal{H}(b) - \mathcal{H}(a) = \mathcal{F}(b) - \mathcal{F}(a). \quad (3.9)$$

Since  $F$  was an arbitrary generalized primitive of  $f$ , we deduce from (3.8) that all primitives of  $f$  are equal up to an additive constant, and therefore  $f$  is integrable in the interval  $G(I)$ . Moreover we have, by virtue of Corollary 2.5, that  $\mathcal{F}(t) = F(G(t)) = \hat{c} + \int_{G(a)}^{G(t)} f(x) dx$  for some  $\hat{c} \in \mathbb{R}$ , and then (3.9) implies (1.2).  $\square$

## 4 Bibliographical reviews on change of variable

(1959) Marcus establishes in [7] sufficient conditions on  $G$  which imply that the composition  $f \circ G$  is integrable. Subsequently, the author gives some results on change of variable for the Riemann integral. Marcus' paper leans on some previous works by Zaslavski on Riemann integrability of compositions which go back to 1953. We drawn all the previous information from Mathematical Reviews.

A remarkable example in [5] shows that, in the conditions of Kestelman's Theorem, the composition  $f \circ G$  needs not be Riemann integrable. Therefore, Kestelman's Theorem generalizes the results in [7].

(1961) Kestelman and Davies publish their respective proofs of Theorem 1.2 in [3, 5]. One can deduce from [3, 5] that both authors were aware of each other's work before the publication of their papers.

We have followed Kestelman's proof in this paper, so the reader already has an idea about the arguments in it. Interested readers might find it useful to know that Lemma 2 in [5] is proven there with some inaccuracies (the set  $S_+$  is expressed as an union of null-measure sets, but not necessarily a countable union, so one cannot deduce that  $S_+$  is null). Lemma 2 in [5] is however correct, as it is a particular case to Theorem 1 in [14]. Anyway, we have benefitted from the proof of [5, Lemma 2], because the ideas in it about Lipschitzian inverses helped us to simplify the arguments we needed from the proof of [14, Theorem 1] and, in particular, we have not used Lebesgue's outer measure.

Modestly, Davies claims in [3] that his paper is "essentially based on the same ideas", which is not true at all. It is based on the following test for the Riemann integrability:  $f(x)$  is integrable over  $[a, b]$  if and only if it is bounded and given any  $\varepsilon, \eta > 0$  there exists a subdivision of  $[a, b]$

such that the intervals in which the oscillation of  $f$  is greater than  $\eta$  have total length less than  $\varepsilon$ . Davies' elementary proof of Theorem 1.2 is really a masterpiece: it simply uses the most basic elements in Riemann integration (namely, Riemann sums) in a wonderfully clever way. It is however very subtle and, therefore, hard to follow or to convey to students.

(1968) Varberg [23] extends Marcus' results by stating and proving the following theorem:

**Theorem 4.1** *Let  $f$  be Riemann integrable on  $[a, b]$  and let  $G$  be absolutely continuous on  $[c, d]$  with  $G([c, d]) \subset [a, b]$  and  $G'$  (up to redefinition on a null-measure set) Riemann integrable on  $[c, d]$ . Then  $(f \circ G) \cdot G'$  is Riemann integrable on  $[c, d]$  and  $\int_{G(c)}^{G(d)} f(x) dx = \int_c^d f(G(t))G'(t) dt$ .*

Note that the assumptions on  $G$  imply that  $G(t) = G(a) + \int_a^t G'(s) ds$  for all  $t \in [c, d]$ , with  $G'$  Riemann integrable. In turn, functions  $G$  in the conditions of Kestelman's Theorem are absolutely continuous (Lipschitz continuous, actually) and  $G' = g$  almost everywhere, so Varberg's Theorem is equivalent to Kestelman's.

**Remark 4.1** *Cleverly, Varberg imposes assumptions over  $f$  on an interval  $[a, b]$  containing the range of  $G$ , and not exactly on the range of  $G$  as we did in our statement of Kestelman's Theorem.*

*Varberg's statement avoids the inconsistency of Riemann integrability of  $f$  on a singleton set, which does arise in our statement in case  $G$  is a constant function.*

*Why didn't we follow Varberg's more rigorous writing?*

*First, the change of variables formula really needs no assumption outside the range of  $G$ , and this should remain clear from the very beginning.*

*Second, in the well-known convention  $\int_a^a f(x) dx = 0$  ( $a \in \mathbb{R}$ ) for the Riemann integral underlies the idea that "all functions are Riemann integrable on singletons". We are convinced that most readers will understand from our statement that no hypothesis on  $f$  is required when  $G$  is constant, and this is exactly what we mean.*

*Finally, the case  $G$  constant is trivial and no one would ever use a theorem to study it. In our opinion, statements should not be overloaded just to avoid a minor abuse of language that only occurs in a trivial situation.*

(1970) Preiss and Uher proved Theorem 1.3 in [12] by adjusting Davies' ideas in [3].

(1981) Navrátil revises Preiss and Uher proof of Theorem 1.3 in [10].

(1985) Terëkhina [19] proves the following result on change of variables:

**Theorem 4.2** *Let  $f$  be Riemann integrable on  $[a, b]$  and let  $G$  have a Riemann integrable derivative on  $[c, d]$ ,  $G([c, d]) \subset [a, b]$ , and  $G(c) = a$  and  $G(d) = b$ . Then  $(f \circ G) \cdot G'$  is Riemann integrable on  $[c, d]$  and  $\int_{G(c)}^{G(d)} f(x) dx = \int_c^d f(G(t))G'(t) dt$ .*

Theorem 4.2 is a particular case to Kestelman's Theorem. Indeed, the Second Fundamental Theorem of Calculus (see [15, page 286]) guarantees that  $G(t) = G(c) + \int_c^t G'(s) ds$  for all  $t \in [c, d]$ .

Terëkhina also introduces in [19] an example showing that  $f \circ G$  needs not be Riemann integrable in the assumptions of Theorem 4.2.

(1996) Sarkhel and Výborný [13] rediscover Theorem 1.3, without being aware of [3, 5, 10, 12]. Their proof differs from those in [3, 10, 12] in the sense that they use Lemma 2.6. In this sense, the proof in [13] is closer to that in [5]. The proof given in section 3 follows essentially the steps in [13], but we avoid the use of the change of variable formula when  $G$  is monotone.

Besides a new proof, we owe to [13] at least two very interesting remarks: first, Theorem 1.3 is better than its analogues for the Lebesgue integral in the sense that we can deduce Riemann integrability of  $f$  from that of  $(f \circ G)g$ , and this is not possible in general with Lebesgue integrals<sup>3</sup>; second, a nice example which shows that we cannot simply omit the assumption that  $f$  be bounded.

(1997) Popovici and Bencze rediscover Kestelman's Theorem in [11] and introduce a new elementary proof, similar to Davies'.<sup>4</sup>

(1998) Cater [2] studies the change of variable formula in terms of the Dini derivatives of  $G$ . The use of Dini derivatives in connection with the formula of change of variable is classical, see [4], but Cater's approach is new and not restricted to monotone substitutions.

For completeness, we state the main result in [2]. Notice that our next statement gathers the information given in Theorem 1 and Corollary 2 in [2]:

---

<sup>3</sup>We know, however, that such a result for the Lebesgue integral is valid when  $G$  is monotone. We are grateful to Professor Rudolf Výborný for having sent us a manuscript of his with a proof.

<sup>4</sup>We are grateful to Professor Mihály Bencze for having sent us the paper [11].

**Theorem 4.3** *Let  $DG$  denote one of the four Dini derivatives of a continuous real valued function  $G : I = [a, b] \rightarrow \mathbb{R}$  (the same Dini derivative at all  $x$ ). Let  $G(a) \leq G(b)$  and let  $DG$  be bounded on  $I$ . Let  $f$  be a bounded function on  $G(I)$  such that for almost every  $t \in [a, b]$ , one or both of the functions  $f \circ G$  or  $DG$  is continuous at  $t$ .*

*In the previous conditions the upper and lower Darboux integrals satisfy the following inequality:*

$$\begin{aligned} \overline{\int_a^b} f(G(t)) DG(t) dt &\geq \overline{\int_{G(a)}^{G(b)}} f(x) dx \\ &\geq \underline{\int_{G(a)}^{G(b)}} f(x) dx \geq \underline{\int_a^b} f(G(t)) DG(t) dt. \end{aligned}$$

*Moreover, if  $(f \circ G) DG$  is Riemann integrable on  $I$ , then  $f$  is Riemann integrable on  $G(I)$  and  $\int_{G(a)}^{G(b)} f(x) dx = \int_a^b f(G(t)) DG(t) dt$ .*

We know, see [4, page 455], that if  $G$  satisfies the conditions of Theorem 1.3 and  $G(a) \leq G(b)$  then  $G$  satisfies the conditions of Theorem 4.3. In this sense, Cater's Theorem is better than Theorem 1.3 as it allows more types of substitutions. The price for that generality turns out to be the assumption on almost everywhere continuity of  $f \circ G$  or  $DG$ , which cannot be omitted, see [2, Example 3].

On the other hand, we note that if  $DG$  in Theorem 4.3 is Riemann integrable then, see [4, page 456],  $G(t) = G(a) + \int_a^t DG(s) ds$  for all  $t \in I$ , so  $G$  satisfies the assumptions in Theorem 1.3 with  $g = DG$ . Examples of functions  $G$  in the conditions of Theorem 4.3 and such that  $DG$  is not Riemann integrable are subtle, see [2, Example 2].

(2001) Bagby wrote in [1] the most recent paper we known about Kestelman's Theorem. Bagby's version applies for functions  $f$  assuming values in an arbitrary Banach space. The proof is elementary, similar to that in [3].

(2008) An extension of Kestelman's Theorem to higher dimensions with injective  $G$  is introduced in [9].

## References

- [1] R. J. Bagby, The substitution theorem for Riemann integrals, *Real Anal. Exchange* **27** (2001/02), 1, 309–314.

- [2] F. S. Cater, A change of variables formula for Darboux integrals, *Real Anal. Exchange* **24** (1998/99), 1, 467–476.
- [3] R. O. Davies, An elementary proof of the theorem on change of variable in Riemann integration, *Math. Gaz.* **45** (1961), 351, 23–25.
- [4] E. W. Hobson, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series V1 (1921)*, Kessinger Publishing, 2008 (Original Edition: Cambridge University Press, 1921).
- [5] H. Kestelman, Change of variable in Riemann integration, *Math. Gaz.* **45** (1961), 351, 17–23.
- [6] J. J. Koliha, Mean, Meaner, and the Meanest Mean Value Theorem, *Amer. Math. Monthly* **116** (2009), 356–361.
- [7] S. Marcus, Sur la superposition de deux fonctions intégrables au sens de Riemann et sur le changement de variable dans l'intégrale de Riemann (Russian), *Rev. Math. Pures Appl.* **4** (1959), 381–389.
- [8] E. J. McShane, *Integration*, Princeton University Press, Princeton, N. J., 1944.
- [9] Z. Molnár, I. Nagy and T. Szilágyi, A change of variables theorem for the multidimensional Riemann integral, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **51** (2008), 121–140 (2009); arXiv:0804.2333v1 [math.CA] 15 Apr 2008.
- [10] J. Navrátil, A note on the theorem on change of variable in a Riemann integral (Czech. English summary), *Časopis Pěst. Mat.* **106** (1981), 1, 79–83.
- [11] F. Popovici and M. Bencze, A new and simple proof of the change of variable formula in the Riemann integral in very general conditions, *Octagon Math. Mag.* **5** (1997), 1, 10–18.
- [12] D. Preiss and J. Uher, A remark on the substitution for the Riemann integral. (Czech. English summary), *Časopis Pěst. Mat.* **95** (1970), 345–347.
- [13] D. N. Sarkhel and R. Výborný, A change of variables theorem for the Riemann integral, *Real Anal. Exchange* **22** (1996/97), 1, 390–395.



- [14] J. Serrin and D. E. Varberg, A general chain rule for derivatives and the change of variables formula for the Lebesgue integral, *Amer. Math. Monthly* **76** (1969), 514–520.
- [15] M. Spivak, *Calculus*, Cambridge University Press, corrected third edition 2006.
- [16] K. R. Stromberg, *An Introduction to Classical Real Analysis*, Wadsworth Inc., Belmont, California, 1981.
- [17] E. Talvila, Characterizing integrals of Riemann integrable functions, *Real Anal. Exchange* **33** (2007), 487–488.
- [18] E. Talvila, The regulated primitive integral, *Illinois J. Math.* **53** (2009), 4, 1187–1219.
- [19] E. Yu. Terëkhina, Change of variable in the Riemann integral (English, Russian original), *Mosc. Univ. Math. Bull.* **40** (1985), 75–79; translation from *Vestn. Mosk. Univ., Ser. I Mat. Mekh.* **103** (1985), 3, 78–80.
- [20] B. S. Thomson, Characterization of an indefinite Riemann integral, *Real Anal. Exchange* **35** (2010), 487–492.
- [21] W. F. Trench, *Introduction to real analysis*, free edition downloaded from <http://ramanujan.math.trinity.edu/wtrench/misc/index.shtml> (previously published by Pearson Education), 2003.
- [22] Ch. J. de la Vallée Poussin, Sur l'intégrale de Lebesgue, *Trans. Amer. Math. Soc.* **16** (1915), 435–501.
- [23] D. E. Varberg, The change of variables formula for Riemann integrals, *Bull. Math. Soc. Sci. Math. R. S. Roumanie (N.S.)* **12** (60) 1968, 4, 239–240 (1969).